HFL/ONR 86-2

(1)

A174960

COLOR CONTRAST METRICS FOR COMPLEX IMAGES

by

David L. Post and Harry L. Snyder

Human Factors Laboratory

Department of Industrial Engineering and Operations Research

Virginia Polytechnic Institute and State University

Blacksburg, VA 24061

September, 1986

SECURITY CLASSIFICATION OF T

REPORT DO

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER

AD-A174 960

ECIPIENT'S CATALOG NUMBER

4. TITLE (and Subtitle)

Color Contrast Metrics for Complex Images

PE OF REPORT & PERIOD COVERED

Technical Report 1982-1983

6. PERFORMING ORG. REPORT NUMBER

HFL/CNR 86-2

7. AUTHOR(s)

David L. Post and Harry L. Snyder

8. CONTRACT OR GRANT NUMBER(s)

N00014-78-C-0238

9. PERFORMING ORGANIZATION NAME AND ADDRESS

Human Factors Laboratory
Virginia Polytechnic Institute and State University
Blacksburg, VA 24061

10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS

NR 196-55

11. CONTROLLING OFFICE NAME AND ADDRESS

Office of Naval Research, Code455
800 N. Quincy STreet
Arlington, VA 22217

12. REPORT DATE

September, 1986

13. NUMBER OF PAGES

90

14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office)

15. SECURITY CLASS (of this report)

Unclassified

15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Visual perception
Color vision
Uniform chromaticity scales
Visual displays
Legibility
CRT displays
Visual detection

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

Designers of color displays have no simple means to relate suprathreshold color contrast to performance because few of the requisite experiments have been performed. One of the obstacles has been the absence of a stanardized and perceptually uniform method for representing color differences. An experiment was performed to explore the utility of several candidate metrics for predicting response speeds to colored numerals superimposed upon spatially complex multicolored backgrounds. The results show that several simple metrics accounted for substantial portions of the sample variation, despite the superficially apparent difficulties of the modelling problem.

Keywords:

DD 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

PAGES _____

ARE

MISSING

IN

ORIGINAL

DOCUMENT

ACKNOWLEDGEMENTS

This research has been supported by the Office of Naval Research, Arlington, Virginia, (Contract N00014-78-C-0238), monitored by Mr. Gerald S. Malecki.

Appreciation is expressed to Messrs. Willard W. Farley and L. Hardy Mason for assistance with hardware and software development, and to Dr. Thomas M. Lippert for assistance with methodological development, hardware calibration, and analysis of his data.

TABLE OF CONTENTS

page

ACKNOWLEDGEMENTS.....i

TABLE OF CONTENTS.....ii

LIST OF TABLES.....v

LIST OF FIGURES.....vii

INTRODUCTION.....1

Purpose of Research.....3

METHOD.....7

Experimental Design.....7

Subjects.....7

Apparatus.....9

Display system.....9

Head-Up display.....11

Background images.....16

Task.....18

Responses.....18

Postural control.....19

Procedure.....20

Scheduling.....20

Sessions.....21

RESULTS.....22

Preliminary Checks.....22

Analysis of Variance.....23

Regressions.....30

Two-Degree Model of Color Contrast.....31

Two-factor ΔE regressions.....32
Four-factor regressions.....34
Comparisons among optimum models.....39
Comparisons with Lippert (1983).....39
Outline Model of Color Contrast.....44
ΔE Regressions.....46
Number of discriminable outline pixels.....47
Proportion of discriminable outline pixels.....48
Summary of results for outline model.....49
Color Variation Metrics.....51
Definitions of variation metrics.....52
Tests of variation metrics.....53
Two-degree Model of Color Variation.....58
Three-factor ΔE regressions.....58
Five-factor regressions.....60
Comparisons among optimum models.....63
Comparisons with Lippert (1983).....63
Nonlinear regressions.....66
DISCUSSION AND CONCLUSIONS.....69
REFERENCES.....74
APPENDIX
1. Average Photometric Measurements for HUD Colors.....76
2. Description of Photographs Used as Backgrounds.....77
3. Characterization of Two-Degree Averages for Backgrounds.....79
4. Chromaticity Plot for Background Images and HUD Colors.....80

5. Colors in the Natural Environment.....81

6. Chromaticity Plot of Colors in the Natural Environment.....83

LIST OF TABLES

TABLE page

1. Analysis of Variance Summary Table.....24
2. Two-Degree Model of Color Contrast: Two-Factor ΔE Regressions.....33
3. Two-Degree Model of Color Contrast: Four-Factor Regressions.....36
4. Correlations Among Squared Distances Along the Color Spaces' Axes.....38
5. Two-Factor ΔE Regressions on Lippert's (1983) Data.....40
6. Four-Factor ΔE Regressions on Lippert's (1983) Data.....41
7. Two-Degree Model of Color Contrast vs. Lippert (1983) Models: Predicting Present Study's Results.....43
8. Two-Degree Model of Color Contrast vs. Lippert (1983) Models: Predicting Lippert's (1983) Results.....45
9. R^2s for Outline Model Regressions.....50
10. R^2s for Color Variation Metrics.....55
11. Correlations Between Color Variation and ΔE in Each Color Space.....56
12. Correlations Between Color Variation Metrics and Squared Distances for Color Spaces Axes.....5
13. Two-Degree Model of Color Variation: Three-Factor ΔE Regressions.....59
14. Two-Degree Model of Color Variation: Five-Factor Regressions.....61

15. Two-Degree Model of Color Variation vs. Lippert (1983) Models: Predicting Present Study's Results.....64

16. Two-Degree Model of Color Variation vs. Lippert (1983) Models: Predicting Lippert's (1983) Results.....65

LIST OF FIGURES

FIGURE page

1. Display configuration for Lippert (1983).....4
2. Block diagram of one replication for one subject.....8
3. Display configuration for hypothetical trial.....13
4. Mean RS for each session.....26
5. Mean RS for each HUD color.....28

INTRODUCTION

Designers of color displays have no simple means to relate suprathreshold color contrast to performance because few of the requisite experiments have been conducted. One of the obstacles has been the absence of a standardized and uniform method for representing color differences. The work which has been performed under the present research contract has addressed this problem by exploring the utility of metrics which represent color differences as distances between colors in a perceptually uniform color space

The initial research in this series compared the 1976 CIE L*u*v*, 1976 CIE L*a*b*, and Cohen and Friden's (1975) Wab theoretically uniform color spaces and produced equations for transforming distance (ΔE) within each space into equivalent achromatic contrast. The experimental method consisted of having subjects represent their perceptions of suprathreshold color contrasts in terms of achromatic luminance contrasts. In an early experiment, subjects adjusted the luminances of seven colors until their brightnesses matched those of 35, 50, and 70-cd/m^2 achromatic standards. This produced a set of 21 standard colors which were used in the subsequent two experiments. In the next experiment, the colors were presented in brightness-matched pairs and subjects adjusted an adjacent, achromatic pair until its luminance contrast matched the

color contrast of the chromatic pair. The subsequent experiment was an expanded replication which involved presenting all possible color pairings rather than only brightness-matched pairs. See Costanza (1981) and Post, Costanza, and Lippert (1982) for details of these experiments.

The results showed that regression models of color contrast accounted for 38% to 80% of the variation in the group's mean achromatic settings. The predictive power of all models was lessened in the second color contrast experiment. Interestingly, though, the ΔE-type models did not yield optimum results for the CIE spaces. Instead, multiple regression equations which individually weighted the distances along each axis provided notably larger values of R^2. These results suggest that the magnitudes of perceived color differences are represented uniformly in Wab space but not in L*u*v* or L*a*b*, at least when they are represented by achromatic contrast. This result is so because the multiple regression model for Wab produced no notable improvement in R^2 over the ΔE_{Wab} model, whereas the L*u*v* and L*a*b* models did show such an improvement. Thus, Wab provided a more uniform model than did the 1976 CIE spaces, but modifying the CIE spaces produced superior predictive power.

During the second phase of the research, emphasis shifted from the perception of color differences to the effects of such differences on human visual task performance. Lippert (1983) performed a study to evaluate the effects of color contrast on accuracy and response speed for reading dot matrix numerals. A head-up display (HUD) was simulated by superimposing a simplified vertical situation display on a static computer-generated background and manipulating color contrast for the HUD's numerals vs. the background. (See Figure 1 for an illustration of the display.) The colors' tristimulus values were known and, therefore, it was possible to specify color contrast in terms of ΔE in the various uniform color spaces. The results of this research were presented by Lippert, Farley, Post, and Snyder (1983) and are contained in more detail in reports by Lippert (1983) and Lippert and Snyder (1986).

Purpose of Research

The study by Lippert (1983) involved the presentation of numerals on large, uniform backgrounds with multicolored surrounds. For many applications, this is a satisfactory model of the viewing conditions. In particular, the results should be applicable to the design of color displays for interactive computer terminals and for computer-generated imagery simulations. However, in some applications, the images are more complex, involving spatially non-uniform,

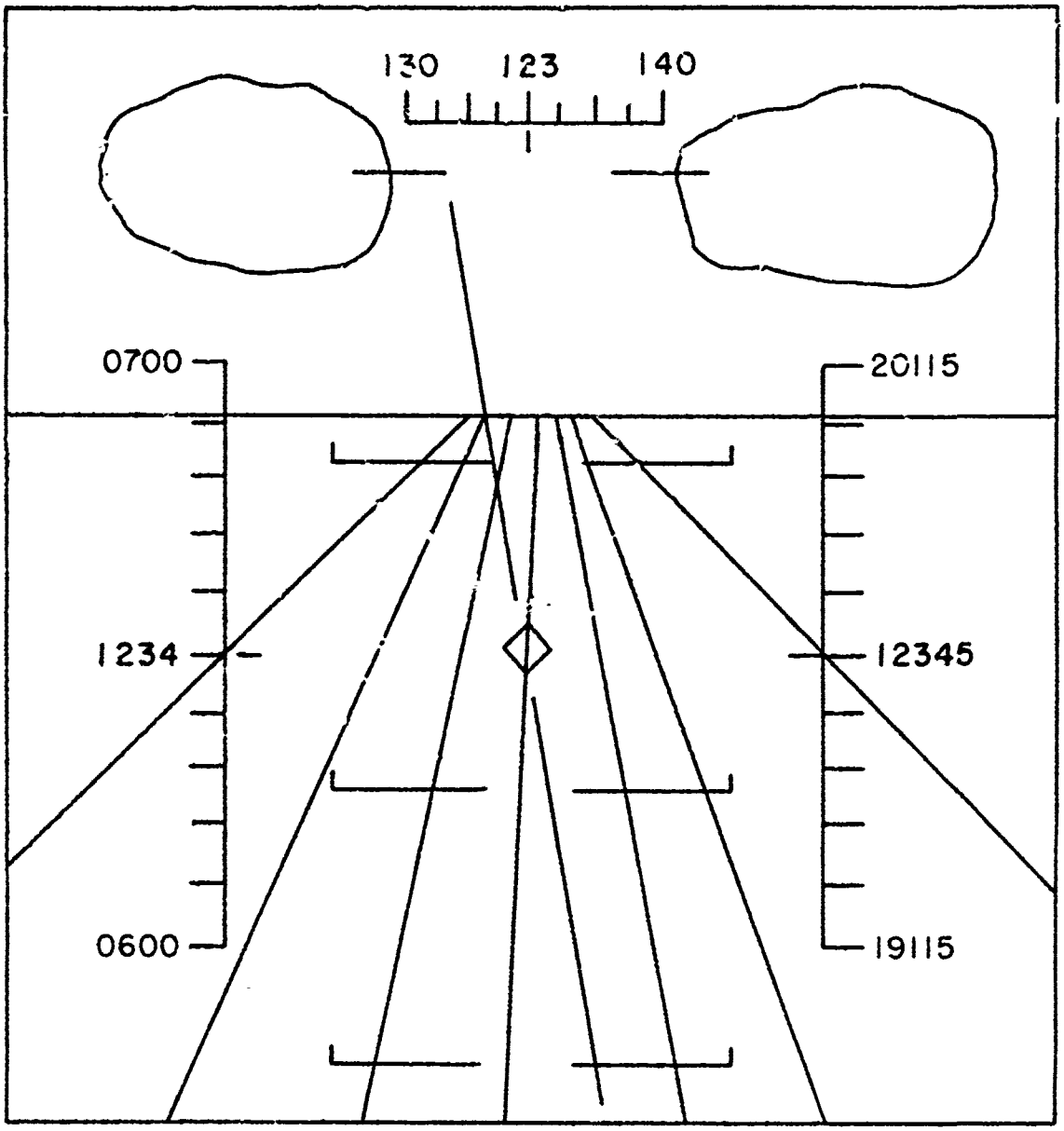


Figure 1. Display configuration for Lippert (1983).

multicolored backgrounds, and may consist of a mixture of computer-generated and real-world imagery.

Head-up displays are a particularly interesting example of this sort of design problem because they combine computer-generated symbology, which the designer can control, with dynamic real-world imagery over which the designer has very little control. In a typical HUD, computer-generated graphics and alphanumerics are projected from an electronic display and collimated. Collimation causes the rays of light from the display to become parallel and, thus, the eye can focus them only by focusing at optical infinity. The illumination is reflected into the viewer's eyes by a combining plate which also transmits illumination from the outside world. Thus, a viewer who is focused at infinity sees an image of the outside world with the computer-generated imagery "watercolored" upon it. The reader has probably seen similar images when looking through windows at night.

Typically, HUDs are used in military aircraft to permit the pilot to see critical flight information without having to look away from the windshield. (Hence, the name "head-up display".) This is particularly advantageous during various phases of landing, weapons delivery, and combat maneuvering and contributes significantly to both the pilot's safety and the probability of completing the mission

requirements. In the near future, it is likely that HUDs will also be used in commercial aircraft and, possibly, automobiles.

It can be seen that a HUD provides a convenient context for studying complex images containing suprathreshold color contrast. The remainder of this document describes an experiment which used a simulated HUD and real-world imagery for this purpose. The goal was to test and extend the generalizability of Lippert's (1983) research by addressing the problem of specifying color contrast for spatially complex multicolored images and relating it to human performance. Models were developed to predict response speed for reading dot matrix numerals superimposed upon photographic backgrounds as a function of color contrast. Comparisons of the results for differing models have implications regarding the geometry and utility of several uniform color spaces and regarding the role of spatial relationships among colors in determining effective color contrast. Such models would provide useful information to designers of color displays and contribute significantly to the modelling of color perception.

METHOD

Experimental Design

Eight subjects viewed a color CRT which displayed digitized color photographs and a superimposed HUD. The color contrast between the HUD and the background was controlled and the subjects performed a task which required reading dot matrix numerals from the HUD. The numerals which were read appeared only in two designated task areas and only one area was read for any given trial. The design was completely within-subjects with four replications per subject, five HUD colors, and 20 backgrounds. Figure 2 illustrates one replication for a given subject. During the course of the experiment, 6400 trials (8 subjects x 4 replications x 2 task areas x 5 HUD colors x 20 backgrounds) were administered. This design yielded 6400 readings and 6400 associated response times (RTs). Afterwards, the data were transferred to the University's IBM 370 for analysis via a standard statistical package.

Subjects

Eight student volunteers (five females) were paid $50 each for participating in the experiment. Seven of the subjects had participated in the previous experiment by Lippert (1983). The use of these subjects seemed advantageous because their experience with a very similar task probably increased the stability of their responses.

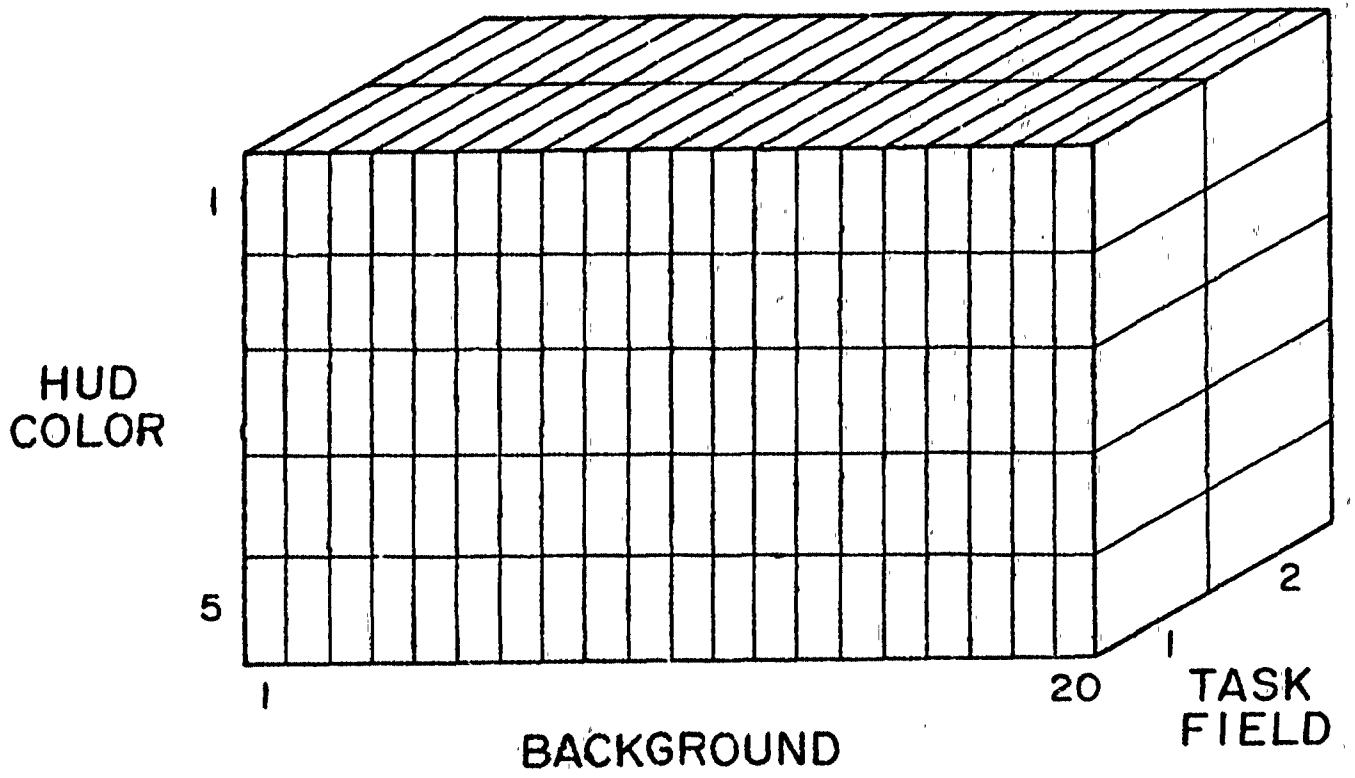


Figure 2. Block diagram of one replication for one subject.

All subjects were screened for normal color vision using the Dvorine Pseudo-Isochromatic Color Plates. The rejection criterion was two misses. The subjects were also tested for 20/20 uncorrected near visual activity using a Bausch and Lomb Orthorater.

Apparatus

Display system. An Aydin Model 8025 19-inch diagonal high-resolution color monitor served as the display. It employs a three-gun (red, green, and blue) Mitsubishi shadow mask CRT incorporating P22 phosphors and a .29 mm triad pitch, i.e., the separation between phosphor triads is .29 mm. The raster is 2:1 positively interlaced and paints a complete image once every 1/30 s. The aspect ratio was adjusted to 1:1, yielding a $(26\ cm)^2$ active display area. The subjects were seated with their eyes approximately 48 cm from the screen. Thus, the display's active area subtended approximately 30 degrees x 30 degrees visually.

The monitor was driven by an International Imaging Systems (IIS) Model 70 digital image processor. This device operated under the control of a Digital Equipment Corporation PDP-11/55 minicomputer and provided 512^2 picture elements ("pixels") at the CRT screen. For a viewing distance of 48 cm, this yielded a limiting spatial resolution of approximately 8.5 cycles/degree. The IIS command voltage for each gun is adjustable through 1024

discrete steps, permitting 1024^3 unique sets of tristimulus values to be generated and specified separately for each pixel.

In order to know the tristimulus values associated with a given triplet of command voltages (or "bit settings"), it is necessary to perform radiometric measurement of the monitor's response characteristics. Characterization is accomplished by measuring separately each gun's spectral radiance distribution at 32 bit settings and computing the associated tristimulus values. Responses at intermediate settings are then predicted via linear interpolation. Note that tristimulus values are additive and that the guns operate independently. Thus, the tristimulus values associated with any combination of bit settings can be predicted from measurements of each gun by itself. Past experience has indicated that characterization remains valid within one just-noticable-difference (JND, i.e., +/- .005 unit for the CIE chromaticity coordinates and +/- 2% for luminance) for approximately two weeks.

The radiometric measurement system consists of several pieces of equipment manufactured by Gamma Scientific, Inc. To measure a light source, some of its emission is collected by a fibre-optics cable and fed to a monochrometer. The monochrometer samples the spectral radiance distribution from 380 to 760 nm in 1-nm increments

under the control of the minicomputer. These narrow bandwidth samples are converted to voltages and amplified by a photomultiplier tube, which is connected to the minicomputer via an analog-to-digital converter. The computer records the samples and computes tristimulus values.

The radiometric measurement system is calibrated by scanning a standard light source whose spectral radiance distribution is known. This distribution is recorded in a computer file and compared with measurements from calibration scans. Any differences between predicted vs. measured values are assumed to be due to error in the measurement system. The comparison yields a correction factor for each 1-nm increment. When a source (e.g., the monitor) having an unknown spectral distribution is scanned, the correction factors are used to produce a set of calibrated radiometric measures.

Head-Up display. The HUD was generated by drawing lines in a portion of the IIS known as the graphics plane. The lines appeared to be continuous but consisted actually of pixel-vectors whose elements were assigned appropriate bit settings. Any pixel in graphics can override pixels at the same x-y coordinates in any other portion of the IIS. This option was used so that the HUD pixels would mask corresponding pixels in the background image, rather than

adding to them. Thus, the HUD's tristimulus values were independent of the background's.

Figure 3 illustrates the appearance of the HUD for a hypothetical trial. The task stimuli consisted of 7 x 9 dot matrix Huddleston font numerals. All digits were changed randomly for each trial. However, experimental tasks were associated only with the centermost digits on the vertical scales. Hereafter, these areas of the HUD will be referred to as fields 1 and 2, respectively. Field 1 always contained four digits while field 2 always contained five digits. For any given trial, the subject responded to only one of the fields. Note that fields 1 and 2 were in "mirrored" positions. Thus, for any given background image, virtually identical color contrast could be produced for either field by reversing the image left to right.

Five HUD digit colors were selected for use in the experiment. The actual chromaticity coordinates and luminances for the colors, averaged over several sets of measurements, are shown in Appendix 1. (As a reference aid, all measurements, etc. pertaining to the experimental stimuli are contained in Appendices 1 - 6.) The first three sets of chromaticity coordinates are very similar to those which were used by Lippert (1983). They were selected because they are representative of those which can be achieved using existing or proposed HUD technology. The

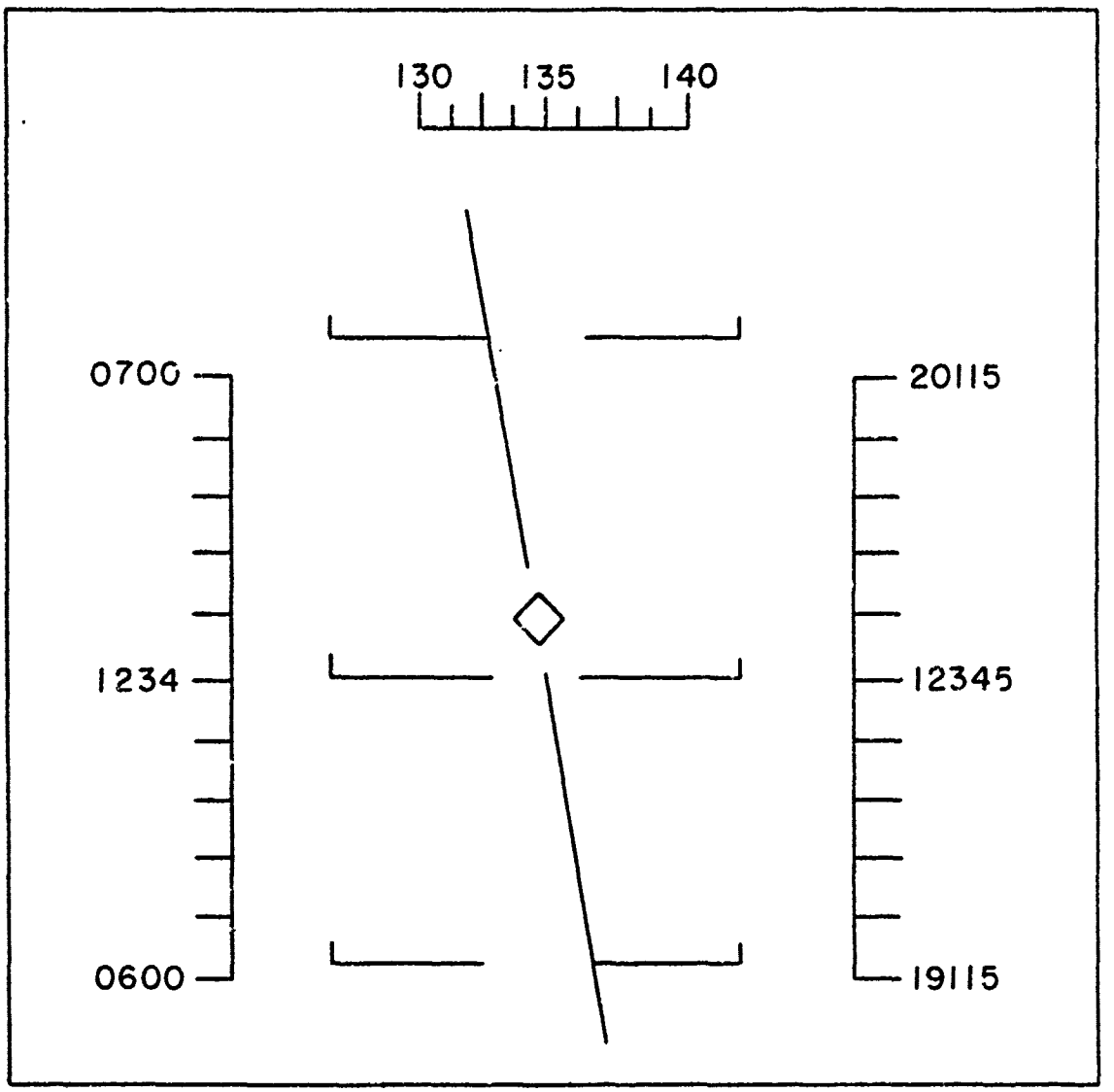


Figure 3. Display configuration for hypothetical trial.

green and blue HUDs were added to increase the number of possible color combinations. The red, green, and blue HUDs were produced using only one of the monitor's guns (i.e., red, green, or blue, respectively) while the achromatic and yellow-green HUDs were generated via appropriate mixtures of the three guns. The average measured luminance for the HUDs was approximately 40 cd/m^2.

Interstimulus display. An achromatic interstimulus display (ISD), consisting of a $(2\ pixel)^2$ checkerboard grating, was used to maintain the subjects' adaptations to photopic levels of illumination and to provide interference with possible afterimages. The ISD was generated by alternately assigning either black or white bit settings to groups of graphics pixels. The ISD's spatial luminance modulation (i.e., $(L_{max} - L_{min}) / (L_{max} + L_{min})$, where L_{max} = the maximum luminance) was unity because L_{min} was zero. The space-weighted average luminance (i.e., $(L_{max} + L_{min}) / 2$) was constant for all trials to avoid confounding the effects of color contrast with variations in the subjects' state of luminance adaptation. The actual luminance level was not particularly critical but was approximately the average of all the HUD/background combinations (HUDBACKs) so as to minimize the likelihood that the sudden changeovers would produce visual discomfort. The ISD's white pixels had approximately the same chromaticity coordinates as the achromatic HUD's.

The IIS design made it possible for the HUD and ISD to coexist in graphics while remaining available for individual display. Thus, it was not necessary to draw these entities each time they were to be presented. (It was necessary, however, to draw new stimuli in the HUD while the ISD was on. This process required approximately 1 s.) Instead, either could be selected via a single command. The time requirement for initiating a changeover was negligible and, thus, the total time requirement was determined by the monitor's refresh rate. This means that changeover required 1/30 s.

The same spatial characteristics of the ISD which masked afterimages also raised thresholds for subsequent stimuli. Thus, the effects of color contrast may have been attenuated somewhat. However, this poses no special problem because any viewing situation imposes a state of spatial adaptation which can be expected to influence color perception. Nunn (1977), for example, has shown that adaptation to monochromatic gratings has influences on subsequent color matching behavior which are related to the gratings' spatial frequencies. Thus, the nature of the ISD's effects on the perception of subsequent stimuli should not have been especially important because its spatial frequency was constant.

Background images. Ten color photographs were selected from a large pool to provide the backgrounds for the experiment. They were chosen on the basis of a subjective evaluation of their color content in the vicinities of the HUD's digits, the objective being to obtain as wide a range of color combinations as possible. By reversing each image left to right, it was possible to use both sides as the background for either task field. Thus, the 10 photographs provided 20 backgrounds which, when combined with the five HUD colors, provided 100 unique color combinations for presentation to the subjects. A brief description of each image is contained in Appendix 2.

Photographs 6 - 8 came from an archive of digital imagery which had been provided to the laboratory several years before by the University of Southern California. The remaining photographs were privately-owned color slides which were digitized using a Nytone Model TSC-1 flying spot scanner, the IIS, and the minicomputer. The digitization process consisted of translating spectral distributions at various points in the slides to bit settings for each CRT gun. This produced a set of 512^2 bit-setting triplets, i.e., 512^2 pixels, for each photograph. The tristimulus values for any of these pixels could be computed afterwards, given the monitor's characterization. Due to the monitor's 1:1 aspect ratio, the images were compressed somewhat in the

horizontal dimension. However, this was not especially important, or even evident, given the nature of the experiment.

The images were stored on a magnetic disk and loaded into the IIS at runtime by the minicomputer, as needed. This permitted the computer to have positive control over their presentation. Each image was loaded while the ISD was on. This process required approximately 5 s.

Presentation of the HUDBACKs to subjects was synchronized with the monitor's raster so that they were always painted starting at the uppermost, leftmost pixel. Changeover to the ISD, however, was not synchronized and started immediately after the subject indicated completion of viewing. Thus, erasure of the HUDBACKs was initiated at random locations on the monitor's screen and required 1/30s.

Synchronization required a variable amount of time because it could not always be achieved on the computer's first attempt and sometimes required several attempts. As a result, there was a variable lag between the moment at which the subject indicated readiness for viewing and the moment at which painting of the next HUDBACK initiated. This variability was probably beneficial to the experiment because pretesting indicated that there was a tendency for subjects to develop an open-loop "HUDBACK-off" response

which was synchronized with the preceding "HUDBACK-on" response. The variable lag tended to discourage this behavior. During the experiment, the lag was usually very brief but occasionally lasted as long as 5 s.

Task

The task for both fields was to report all digits, including any leading zeroes which were present. The digits were selected randomly from a uniform distribution ranging from 0000 to 9999 in the case of field 1, and ranging from 00000 to 99999 in the case of field 2. The subjects were instructed to minimize their RTs while maintaining complete reading accuracy. This instruction was designed to fix the subjects' speed/accuracy tradeoffs in the maximally conservative mode, thereby reducing unexplained variability within subjects and, possibly, simplifying the subsequent analyses by forcing all of the task-induced variability into their RTs. The subjects were also instructed to focus their eyes in the vicinity of the upcoming task field before triggering each trial. This instruction was intended to eliminate variability within subjects associated with visual search times.

Responses. The subjects were interfaced with the minicomputer via a solid-state microswitch which was equipped with a power supply and connected to the analog-to-digital converter. When the switch closed, the computer

started a trial by initiating changeover to the next HUDBACK and starting a clock. When the switch opened, the computer terminated the trial by stopping the clock and initiating changeover to the ISD. The clock ran at 1 KHz, thereby providing 1-ms resolution for measurement of the subjects' RTs.

After each trial, the subject reported the contents of the appropriate task field to the experimenter, who relayed this information to the computer via its console. The computer recorded the RT and response in a file and set up for the next trial. The RT was corrected so as to account for the delay between the time at which the clock was started and the time at which the first stimulus pixel was painted. After setup was complete, the computer reported the task field for the upcoming trial to the experimenter, who relayed this to the subject. The computer waited for the experimenter to signal readiness to begin and then verified that the switch was open before rendering it operative for initiating the trial.

Postural control. A padded forehead rest was used to control the subjects' viewing distance and level relative to the CRT screen. The chair was a padded secretarial type and its height was adjusted for each subject so that the forehead rest was in a comfortable position. Also, the backrest was adjusted to provide proper support. The wheels

were removed from the chair to prevent it from creeping. The subjects were permitted to pause, stand, stretch, etc., as needed between trials.

The response switch was mounted on a mobile base which rested on a tabletop in front of the monitor. The base consisted of a Dan Wesson combat pistol grip, mounted horizontally on a plywood board which was .635 cm thick. The grip had grooves for the fingers to assure a consistent hand-position and was suitable for left- or right-handed subjects. The subjects used their dominant hand at all times. These precautions were intended to help to stabilize RT. The switch was countersunk into the grip and was positioned for index-finger triggering. The board was large enough to permit the entire hand to rest upon it. This, combined with the board's thinness and the grip's angle, seemed to provide optimum comfort. Note also that this design permitted the subjects to choose from and change among a wide variety of arm-positions while providing full postural support.

Procedure

Scheduling. The subjects were screened for vision and signed an informed consent form before reporting for the first experimental session. Each subject performed one session daily for five consecutive days. Each session constituted one complete replication and required an average

of 45 minutes.

Sessions. For each session, the subjects viewed the ISD for one minute before starting the task, to stabilize their visual adaptation. Each session consisted of 200 trials in which each HUD color was presented with each of the backgrounds for both task fields. All trials for a given HUD color were performed sequentially. The order of presentation for the HUD colors and task fields was randomized within each session. The order of presentation for the backgrounds was randomized within each HUD color.

The first session was a practice session only. Its purpose was to familiarize the subjects with the task and stabilize their behavior. Therefore, the resulting data were not analysed. However, the subjects were not informed of this lest they take the session less seriously and thereby defeat its purpose. The practice session started by having the subjects read an instruction sheet followed by a brief question and answer period and a preview of the background images. Pretesting without the preview had shown that subjects were unable to describe the contents of any of the images after an experimental sequence and were understandably curious about them. Thus, the initial preview reduced the probability that subjects would extend their RTs in order to examine the images. At the end of the final session, each subject was paid and debriefed.

RESULTS

Preliminary Checks

A histogram was plotted which showed the residuals produced by subtracting the appropriate N x HUD color x background mean (where N is the number of digits presented) from each of the 6400 RTs. It contained a pronounced positive skew, as is typical for RT data. Therefore, the data were transformed by taking their inverses. This improved the distribution's symmetry and yielded a dependent measure which can be conceptualized as response speed (RS). The units for RS were, thus, responses/ms.

The proportion of correct responses (PC) was computed by totalling the number of correct responses within each cell for each subject across the four sessions and dividing by four. This yielded 1600 discrete random variables, capable of taking on five different values. A histogram was plotted and contained an extreme negative skew. Approximately 90% of the proportions were 1, 8% were .75, 1.5% were .5, .498% were .25, and the remaining .002% were 0. The grand mean was .963 correct responses. These findings made it clear that PC was insensitive and that the subjects had been successful in following their speed/accuracy instructions. Therefore, no further analysis of PC was attempted.

Analysis of Variance

To verify the presence of significant color-related effects, a four-way, fixed-effects, full-factorial, within-subjects analysis of variance (ANOVA) was performed on the 6400 RS observations. The main effects for this analysis were replications, N, HUD color, and background. The results are shown in Table 1. All effects other than interactions involving replications were significant, $p < .02$ in all cases, and accounted for a total of 46.12% of the sample variation.

Post-hoc comparisons were performed for the main effects using a two-step procedure. Non-significant comparisons were identified via the minimally-conservative LSD test. Significant comparisons were identified via the maximally-conservative Scheffe test. Comparisons which showed significance for the LSD but failed for the Scheffe were regarded as indeterminate. All comparisons were tested using alpha = .05.

The main effect of replications accounted for only .42% of the sample variation, indicating that the subjects' variability over time was relatively low. The means for each session are plotted in Figure 4. Examination of Figure 4 suggests that RS increased gradually at first and then stabilized during the last two sessions. The LSD showed no significant differences between sessions 1 vs. 2, 2 vs. 4,

Table 1. Analysis of Variance Summary Table

Source	df	MS	F	p
Between Subjects (S)	7	.00015873		
Within Subjects				
Replications (R)	3	.00000460	4.30	= .0164
R x S	21	.00000107		
Number of digits (N)	1	.00012063	94.82	< .0001
N x S	7	.00000127		
HUD color (H)	4	.00004917	27.83	< .0001
H x S	28	.00000177		
Background (B)	19	.00003997	54.02	< .0001
B x S	133	.00000074		
R x N	3	.00000015	1.44	= .2583
R x N x S	21	.00000010		
R x H	12	.00000024	.85	= .6044
R x H x S	84	.00000028		
R x B	57	.00000008	1.12	= .2729
R x B x S	399	.00000007		
N x H	4	.00000124		= .0004
N x H x S	28	.00000017		
N x B	19	.00000065	9.91	< .0001
N x B x S	133	.00000007		

Table 1 (continued)

H x B	76	.00000493	28.47	< .0001
H x B x S	532	.00000017		
R x N x H	12	.00000008	1.31	= .2267
R x N x H x S	84	.00000006		
R x N x B	57	.00000006	.88	= .7130
R x N x B x S	399	.00000006		
R x H x B	228	.00000007	1.07	= .2262
R x H x B x S	1596	.00000006		
N x H x B	76	.00000024	3.31	< .0001
N x H x B x S	532	.00000007		
R x N x H x B	228	.00000006	.92	= .7867
R x N x H x B x S	1596	.00000006		

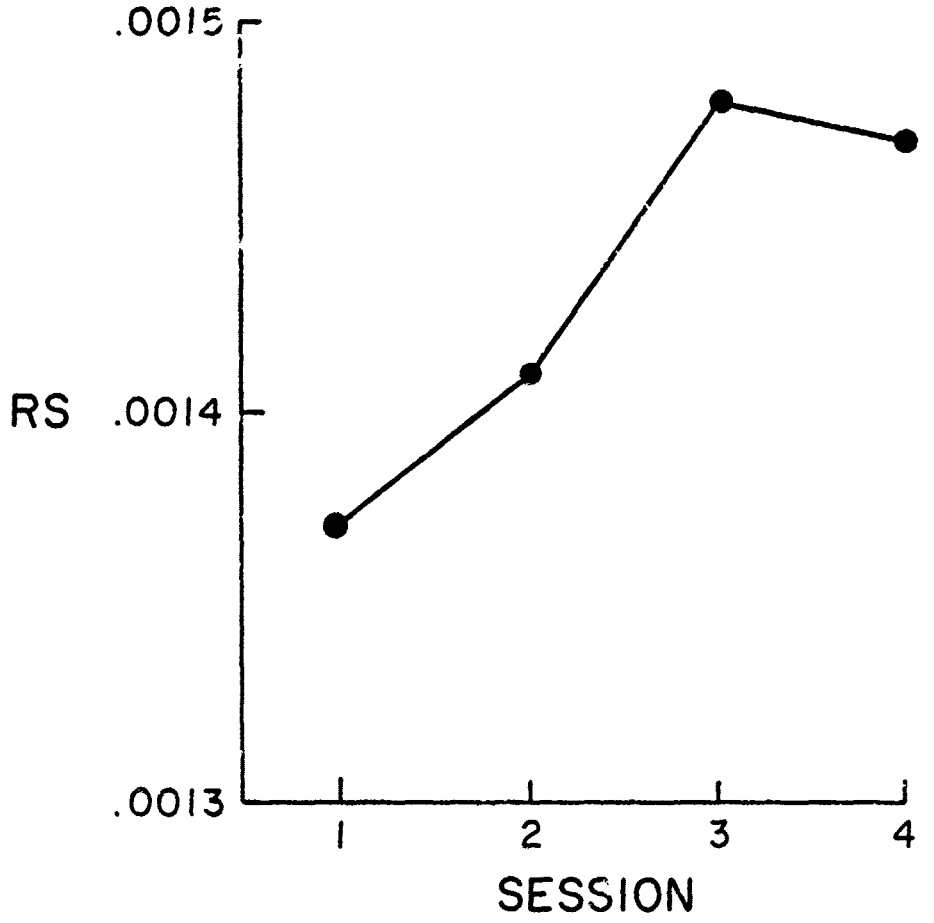


Figure 4. Mean RS for each session.

or 3 vs. 4, while the Scheffe showed session 1 < 3. Thus, there appears to have been a significant improvement in RS over time. The absence of significant interactions involving replications indicates that this improvement was distributed evenly throughout the experimental conditions.

The main effect of N accounted for 3.71% of the sample variation. The mean RSs for the four- and five-digit tasks were .001441 and .001297, respectively.

The main effects of HUD color and background and the HUD color x background interaction accounted for 6.04%, 23.34% and 11.51%, respectively, of the sample variation. Thus, color and color contrast effects accounted for 88.7% of the variation attributable to significant effects. The RS means for each HUD color are charted in Figure 5. The LSD showed no significant differences between achromatic vs. yellow-green and yellow-green vs. green. The Scheffe showed that RS for the red and blue HUDs was significantly greater than for the other three. The main effect of background was broken into seven nonsignificantly-different groupings by the LSD and into five groupings by the Scheffe, showing that many of the backgrounds did not produce statistically unique effects. The HUD color x background interaction is not readily interpretable because it represents color contrast effects. These are analysed subsequently via regression techniques.

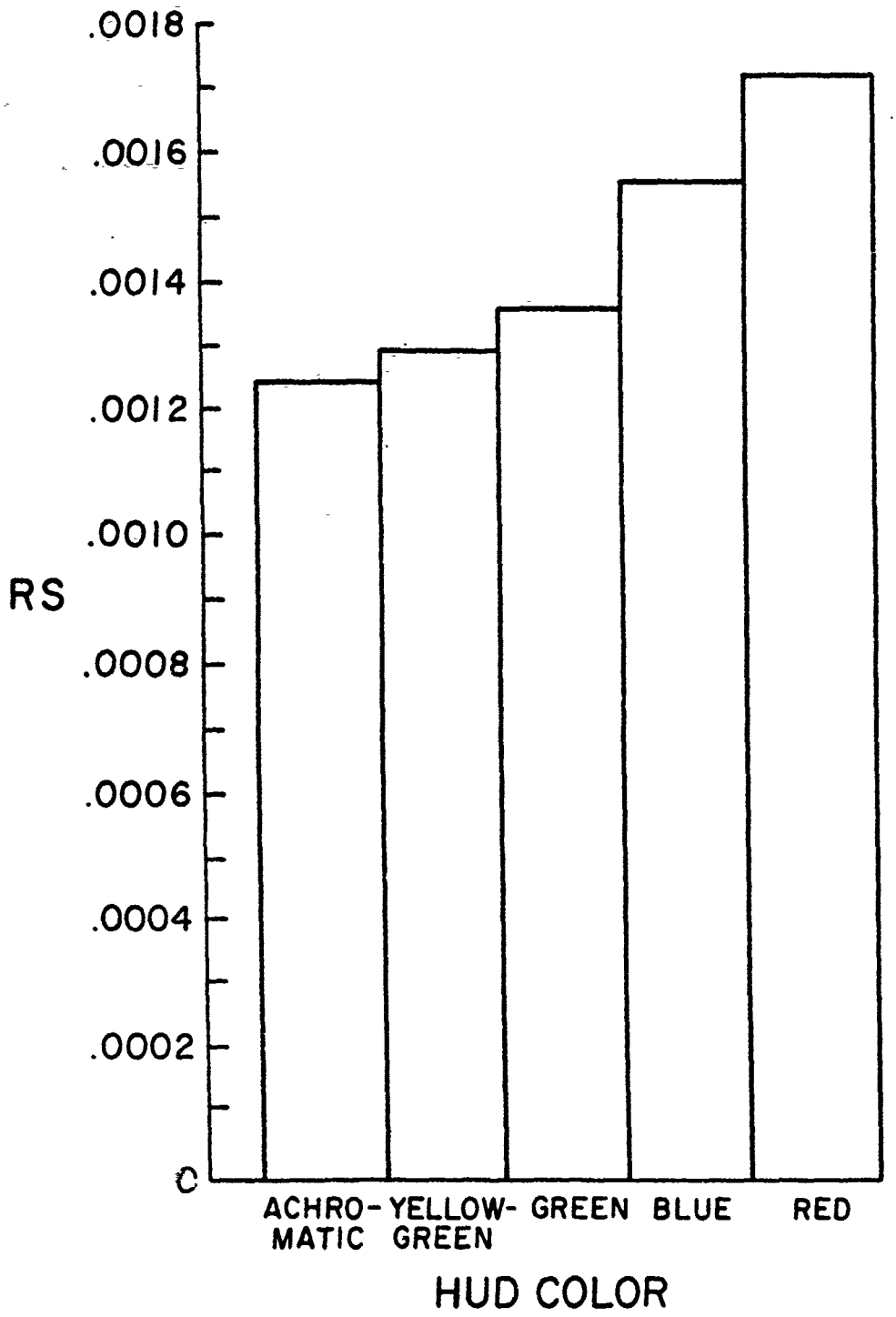


Figure 5. Mean RS for each HUD color.

Although the N x HUD color, N x background, and N x HUD color x background interactions accounted for only .15%, .38%, and .56% of the sample variation, respectively, their statistical significance is instructive. The N x HUD color interaction indicates that there was a difference between the left and right sides of the monitor which varied with HUD color. This probably reflects a minor nonuniformity in convergence. The N x background interaction indicates that the effect of the background varied with the number of digits displayed. This can be taken as evidence that unique contrasts which significantly affected RS occurred for the fifth digit. The N x HUD color x background interaction probably represents the mediating effect of HUD color on the preceding interaction, although other interpretations are possible.

Perhaps the most interesting ANOVA results are those concerning the effect of HUD color on RS. The recommendations of Rizy (1967) and Meister and Sullivan (1969) lead one to expect a significant difference between the red vs. blue HUDs and little difference between blue vs. green. The present findings contradict these expectations. Clearly, recommendations which are intended to apply to achromatic backgrounds did not generalize to a case involving colored backgrounds. Instead, the ordering of the HUD colors as a function of decreasing RS is very similar to

that which was obtained as a function of increasing visual search time in an experiment by Bloomfield (1979), using colored targets and backgrounds. To the extent that the present experiment provided color contrasts which are representative of those encountered with airborne HUDs, the present findings regarding the merits of specific HUD chromaticities may be generalized to the applied setting. Thus, more highly saturated colors yield better performance against a variety of real-world background colors and textures.

Regressions

Although the ANOVA produced some useful information, it did not provide a convenient means for developing metrics which would relate the HUDBACKs' colors to the subjects' performance. The main purpose of the present experiment was to explore the utility of metrics which are based upon distance in the L*u*v*, L*a*b*, and Wab color spaces. Since the color space distance is a continuous variable, this was best approached via regression techniques.

All of the regression analyses model the HUDBACKs' pixels as points within the various color spaces and relate the distances among these points to the subjects' RS. Because of the multicolored nature of the backgrounds, the representation of any particular HUDBACK within a given color space consists of a single point depicting the HUD

color and a cluster of points depicting the background pixels' colors. It would not have been practical to model the color contrast between a given HUD and background using the distances between the HUD and each point in the background cluster. This would have yielded a regression equation containing chousands of parameters, none of which, taken individually, would be expected to have any predictive utility. Furthermore, many of these distances would have pertained to background pixels which were physically distant from the HUD's digits and, therefore, would have had little or no predictive utility under any circumstances. Thus, the first stage of the regression analysis sought to derive a representative of the color-distances between the HUD and background pixels which would use as few parameters as possible and exclude superfluous background pixels. In other words, a global representation of color contrast in the immediate vicinity of the HUD's digits was desired.

Two-Degree Model of Color Contrast

The first approach to modelling global color contrast involved averaging over all background pixels within a two-degree radius (subtended visually) of each digit-set's centermost point. Each of these pixels' 1931 CIE tristimulus values were computed and transformed into each of the uniform color spaces. The results were averaged and the resulting sets of coordinates were taken to represent

the "average" color of the background in each color space. The color contrast between any given HUD and background was then represented as ΔE between the HUD and the two-degree average. As a consequence of the averaging procedure, this representation of color contrast is independent of the particular digits which appeared for any given trial. Thus, color contrast was modelled as if it had been replicated for each subject and session. The rationale for choosing a two-degree radius was that the 1931 CIE standard colormetric observer is recommended only for stimuli which subtend four degrees visually or less.

Two-factor ΔE regressions. The data were averaged over subjects and sessions and a series of two-factor regressions was performed on the RS means. The predictors in the equations were N and ΔE in each color space. A regression using ΔE in 1931 CIE tristimulus space (Tri) was included as a control. The results are shown in Table 2.

The regression coefficients for all parameters in all models were significant, $p < .01$ in all cases. As expected, increasing N was associated with decreasing RS (this had already been demonstrated in the ANOVA) while ΔE showed the opposite relationship. Interestingly, there were no practical differences among the R^2s for the various color spaces, nor were any particularly impressive. Thus, the uniform color spaces failed to distinguish themselves

Table 2. Two-Degree Model of Color Contrast:

Two-Factor ΔE Regressions

Parameter	Regression Coefficient	p	R^2
Intercept	2.294190×10^{-3}	< .01	.261
N	-2.745800×10^{-4}	< .01	
ΔE_{Tri}	6.835931×10^{-6}	< .01	
Intercept	2.124860×10^{-3}	< .01	.271
N	-2.745800×10^{-4}	< .01	
Δ_{L*u*v*}	5.778463×10^{-6}	< .01	
Intercept	2.082100×10^{-3}	< .01	.284
N	-2.745800×10^{-4}	< .01	
ΔE_{L*a*b*}	7.411417×10^{-6}	< .01	
Intercept	2.231760×10^{-3}	< .01	.292
N	-2.745800×10^{-4}	< .01	
ΔE_{Wab}	3.313904×10^{-5}	< .01	

significantly from the nonuniform space upon which they were all intended to improve (i.e., Tri), although all differences were in the expected direction.

Four-factor regressions. Previous experience (see Costanza, 1981; Lippert, 1983; Post et al., 1982) suggested that the R^2s might be improved if the individual contributions of distance along each axis of the three-space were kept separate in the regression equations. This was accomplished via four-factor second-order regressions, e.g.:

$$x_0 + x_1\Delta L^{*2} + x_2\Delta u^{*2} + x_3\Delta v^{*2} = (RS - x_4N)^2 \qquad (1)$$

where x_0 is the intercept, x_1, x_2, and x_3 are the regression slopes for the squared distances along each of the L*u*v* axes, and x_4 is the regression slope associated with N, obtained via the two-factor ΔE regressions. Note that these regressions caused each axis to undergo the optimum linear rescaling for a Pythagorean representation of distance. Of course, the resulting slope for each axis applies to squared distance and, therefore, if one wishes to rescale the axes before computing the squared distances, the appropriate scaling factors are the square roots of the slopes.

To use Equation 1 for predictive purposes, the terms would be rearranged as follows:

$$(x_0 + x_1\Delta L^{*2} + x_2\Delta u^{*2} + x_3\Delta v^{*2})^{0.5} + x_4N = R\ S. \qquad (2)$$

Equations of this form were used to compute the R^2 for the four-factor models.

Movement of N to the right side of the regression equations was necessary to obtain a solution using linear regression methods. The procedure of estimating a slope and then using the estimate as a constant within another regression is unusual but was permissible in this case because N is orthogonal to all other parameters. Thus, the estimate of its slope is independent of all other estimates and vice versa. This is illustrated in Table 2, wherein N's slope is invariant across color spaces. Similar techniques are used sometimes to remove seasonal trends from time series data (see Cook and Campbell, 1979).

The results for the four-factor regressions are shown in Table 3. None of the slopes for Tri's axes differed significally from zero, $p > .13$ in all cases, and each regression for the other three spaces contained one axis whose slope was not significant, $p > .09$ in all cases. Since the distances along the axes had been chosen randomly, it seemed plausible that the nonsignificant slopes might have resulted from covariance among the parameters. Therefore, Pearson product-moment correlations were computed among the squared distances along the axes for each color space. The results, shown in Table 4, confirmed the suspicions regarding covariance. The squared distances for

Table 3. Two-Degree Model of Color Contrast:

Four-Factor Regressions

Parameter	Regression coefficient	p	R^2
Intercept	6.637153×10^{-06}	< .01	.207
N	$-2.745800 \times 10^{-04}$	< .01	
$(\Delta X)^2$	2.231997×10^{-10}	= .13	
$(\Delta Y)^2$	1.936699×10^{-10}	= .19	
$(\Delta Z)^2$	1.087973×10^{-10}	= .43	
Intercept	5.563464×10^{-06}	< .01	.480
N	$-2.745800 \times 10^{-04}$	< .01	
$(\Delta L^*)^2$	9.750373×10^{-10}	< .01	
$(\Delta u^*)^2$	1.173016×10^{-10}	< .01	
$(\Delta v^*)^2$	2.073357×10^{-10}	= .95	
Intercept	5.350179×10^{-06}	< .01	.496
N	$-2.745800 \times 10^{-04}$	< .01	
$(\Delta L^*)^2$	1.001560×10^{-09}	< .01	
$(\Delta a^*)^2$	2.774472×10^{-10}	< .01	
$(\Delta b^*)^2$	$-5.399573 \times 10^{-11}$	= .18	
Intercept	6.078701×10^{-06}	< .01	.265
N	$-2.745800 \times 10^{-04}$	< .01	
$(\Delta W)^2$	3.918890×10^{-09}	< .01	
$(\Delta a)^2$	4.255722×10^{-09}	= .09	
$(\Delta b)^2$	1.388641×10^{-08}	< .01	

all three of Tri's axes were highly correlated and all other spaces produced one or more rs which differed appreciably from zero. The differences among the color spaces for these correlations are especially interesting when one considers that the spaces are all transformations of one another. Note that the statistical significance of the correlations is irrelevant. The fact that they are nonzero proves that the squared distances used in the present experiment are not independent of one another.

Comparisons of the R^2s for the two- vs. four-factor models show that rescaling the axes caused the R^2s for L*u*v* and L*a*b* to increase while the R^2 for Tri and Wab declined (Tables 2 and 3). The significance of these differences was tested using a procedure described by McNemar (1955). This test showed that the differences for L*u*v*, L*a*b*, and Tri were significant, $p < .01$ in all cases, but the comparison for Wab was not, $p = .14$. The reductions for Wab and Tri make sense only if one assumes that they resulted from rounding and/or truncation in the computations.

It is clear that the four-factor models produced the best results for L*u*v* and L*a*b*, indicating that a simple linear rescaling of their axes significantly improves their perceptual uniformity. As for Wab and Tri, the linear rescaling had no benefits and, therefore, the computationally-similar ΔE models emerged as optimum for

Table 4. Correlations Among Squared Distances Along the Color Spaces' Axes

Tri:		
	$(\Delta Y)^2$	$(\Delta Z)^2$
$(\Delta x)^2$	.9802	.8593
$(\Delta Y)^2$		.9117
L*u*v*:		
	$(\Delta u^*)^2$	$(\Delta v^*)^2$
$(\Delta L^*)^2$	- .1755	-.1667
$(\Delta u^*)^2$		.5456
L*a*b*:		
	$(\Delta a^*)^2$	$(\Delta b^*)^2$
$(\Delta L^*)^2$	-.2757	-.1608
$(\Delta a^*)^2$		.0348
Wab:		
	$(\Delta a)^2$	$(\Delta b)^2$
$(\Delta W)^2$	.0290	.6223
$(\Delta a)^2$		.2372

these two spaces. With the exception of the results for Tri, which had not been tested previously, these findings are identical with those reported by Costanza (1981).

Comparisons among optimum models. Although a subjective evaluation of the R^2s for the optimum models was revealing, it was not certain whether the differences among these statistics were significant. The procedure from McNemar (1955) showed that the comparisons between L*u*v* vs. L*a*b* and Wab vs. Tri were not, $p > .17$ in both cases, while the comparisons of L*u*v* and L*a*b* vs. Wab and Tri were significant, $p < .01$ in all cases. Thus, the L*u*v* and L*a*b* spaces outperformed the basic 1931 CIE space when their axes were rescaled but Wab did not under any circumstances.

Comparisons with Lippert (1983). Given the similarities betwen Lippert's (1983) study and the present research, it seemed worthwhile to compare the results from the two experiments. Lippert's (1983) data provide an excellent means for evaluating the present results because differences can be taken to reflect the effects of uniform vs. nonuniform backgrounds. To facilitate the present and following discussions, the results from two-factor ΔE and four-factor regressions on Lippert's (1983) data are summarized in Tables 5 and 6, respectively.

Table 5. Two-Factor ΔE Regressions on Lippert's (1983) Data

Parameter	Regression coefficient	p	R^2
Intercept	3.273270×10^{-3}	< .01	.267
N	-4.065800×10^{-4}	< .01	
ΔE_{Tri}	6.332622×10^{-6}	< .01	
Intercept	3.174630×10^{-3}	< .01	.394
N	-4.065800×10^{-4}	< .01	
ΔE_{Luv}	3.333600×10^{-6}	< .01	
Intercept	3.155000×10^{-3}	< .01	.350
N	-4.065800×10^{-4}	< .01	
ΔE_{Lab}	5.487360×10^{-6}	< .01	
Intercept	3.014483×10^{-3}	< .01	.425
N	-4.065800×10^{-4}	< .01	
ΔE_{Wab}	3.059915×10^{-5}	< .01	

Table 6. Four-Factor ΔE Regressions on Lippert's (1983) Data

Parameter	Regression coefficient	p	R^2
Intercept	1.044046×10^{-05}	< .01	.597
N	$-4.065800 \times 10^{-04}$	< .01	
$(\Delta X)^2$	9.935955×10^{-10}	< .01	
$(\Delta Y)^2$	8.645602×10^{-09}	< .01	
$(\Delta Z)^2$	1.081371×10^{-10}	< .01	
Intercept	1.032066×10^{-05}	< .01	.656
N	$-4.065800 \times 10^{-04}$	< .01	
$(\Delta L^*)^2$	1.470673×10^{-08}	< .01	
$(\Delta u^*)^2$	8.624694×10^{-11}	< .01	
$(\Delta v^*)^2$	4.182645×10^{-11}	< .01	
Intercept	1.003311×10^{-05}	< .01	.619
N	$-4.065800 \times 10^{-04}$	< .01	
$(\Delta L^*)^2$	1.561245×10^{-08}	< .01	
$(\Delta a^*)^2$	2.658661×10^{-10}	< .01	
$(\Delta b^*)^2$	1.811760×10^{-10}	< .01	
Intercept	1.193174×10^{-05}	< .01	.422
N	$-4.065800 \times 10^{-04}$	< .01	
$(\Delta W)^2$	1.510877×10^{-08}	< .01	
$(\Delta a)^2$	5.601355×10^{-09}	< .01	
$(\Delta b)^2$	$-1.096674 \times 10^{-09}$	= .67	

For Lippert's (1983) data, the four-factor models were best for Tri, L*u*v*, and L*a*b* while the two-factor ΔE model was best for Wab. Comparisons among the R^2s for the four-factor models showed that Wab's R^2 was significantly lower than all others, $p < .01$ in all cases, L*u*v*'s R^2 was significantly greater than all others, $p < .01$ in all cases, and the R^2s for Tri vs. L*a*b* did not differ significantly, $p = .15$. Thus, a linear rescaling of the 1931 CIE space produced results which were notably better than Wab's and indistinguishable from L*a*b*'s.

To assist the comparison of the two experiments, the optimum models obtained from analysis of Lippert's (1983) data were used to predict the results from the present study. The R^2s were then compared with those from the present study's optimum models. The procedure from McNemar (1955) showed that the R^2s from L*u*v*, L*a*b*, and Tri were significantly lower using Lippert's (1983) coefficients, $p < .01$ in all cases, but the comparison from Wab was not significant, $p = .15$. These results are summarized in Table 7.

Next, the present study's optimum models were used to predict Lippert's (1983) RS means. The resulting R^2s were significantly lower, $p < .05$ in all cases, although the actual magnitude of the difference for Wab is trivial. These results are summarized in Table 8. All of the tests

Table 7. Two-Degree Model of Color Contrast vs. Lippert (1983) Models: Predicting Present Study's Results

Color space	L model	R^2	P model	R^2	p
Tri	4F	.167	2F-ΔE	.261	< .01
L*u*v*	4F	.379	4F	.480	< .01
L*a*b*	4F	.382	4F	.496	< .01
Wab	2F-ΔE	.277	2F-ΔE	.292	= .15

L = Lippert (1983)

P = Present study

4F = Four-factor

2F-ΔE = Two-Factor ΔE

were performed one-tailed because, for a given color space, it would have been mathematically impossible for the R^2 produced using one experiment's optimum model to predict that experiment's results to be exceeded by the R^2s produced using the other experiment's optimum models.

It is clear that Wab fared well in the cross-experimental comparisons. For the other spaces, though, the attempts to predict across experiments produced substantial decrements in R^2. This suggests that differences in the uniformity of the backgrounds used in the two experiments exerted an important influence on the spaces' coefficients, although the results can also be attributed to subject and color-related differences. More importantly, though, the cross experimental findings show that neither experiment's models of color contrast adequately predicted the other's results. Even though the differences for Wab were slight and/or nonsignificant, the resulting R^2s did not compare favorably with those produced by the optimum models within each experiment.

Outline Model of Color Contrast

The results for the two-degree model were encouraging but comparison of the R^2s with those from Lippert (1983) suggested that a more sophisticated model of the background colors might yield worthwhile improvements. The average R^2 for the optimum models in Tables 5 and 6 is .574 whereas the

Table 8. Two-Degree Model of Color Contrast vs. Lippert (1983) Models: Predicting Lippert's (1983) Results

Color space	L model	R^2	P model	R^2	p
Tri	4F	.597	2F-ΔE	.253	< .01
L*u*v*	4F	.656	4F	.397	< .01
L*a*b*	4F	.619	4F	.295	< .01
Wab	2F-ΔE	.425	2F-ΔE	.406	= .05

L = Lippert (1983)

P = Present study

4F = Four-factor

2F-ΔE = Two-factor ΔE

average for the optimum models in Tables 2 and 3 is .382. Since Lippert's (1983) backgrounds were spatially uniform, one might, therefore, expect that an appropriate model of the present study's backgrounds would yield an average improvement in R^2 of approximately .192, the difference between .574 and .382.

An attempt was made to refine the two-degree model by restricting consideration to background pixels which were immediately adjacent to, or outlined, the digits' pixels. For each of the 6400 observations, the average color coordinates for the outline pixels were computed in each color space for each task-related digit. This permitted the representation of the unique color contrast between each digit and its immediate background for every trial. This approach seemed to offer greater precision than the two-degree model because it considered only those background pixels which might be expected to have the greatest impact on perceived color contrast for the digits and because it represented the unique color contrast resulting from superimposing a particular digit on a given background. Thus, unlike the two-degree model, the outline model treated each trial as if it involved the presentation of a unique color contrast.

<u>ΔE regressions.</u> The first use of the outline model related ΔE in each color space directly to the subjects' RS.

First, the color contrasts were averaged across digits for each trial and the averages and N were regressed on the 6400 observations using two-factor ΔE models. Next, four-and five-factor ΔE regressions were performed for the four- and five-digit task fields, respectively, in which ΔE for each digit was represented separately.

For both sets of regressions, the RS data were pseudo-normalized prior to analysis. For the two-factor ΔE models, each subject's mean RS was subtracted from that subject's observations. This eliminated basic differences among subjects, which the regression equations did not account for. For the four- and five-factor ΔE regressions, each subject's mean RS for each task-field was subtracted because these equations could not include N as a predictor. Thus, for both sets of regressions, the dependent measure became ΔRS.

Number of discriminable outline pixels. The second use of the outline model represented the color contrast for each digit as a function of the number of outline pixels whose color differed by at least one JND from the digit's. In the case of the L*u*v* and L*a*b* spaces, a distance of 1.0 is intended to be equivalent to one JND. For Wab space, there is no known relationship between distance and JNDs, so a distance of 1.0 was selected arbitrarily to represent one JND. For 1931 CIE tristimulus space, JNDs are known to vary

in size and shape but a size of 1.0 was assumed again for the sake of comparison.

For a given color space, a discriminable outline pixel was defined as one whose color lay one unit or more from the digit's color. Two sets of regression analyses, similar to those for the ΔE models mentioned above, were performed using the number of discriminable outline pixels (NDOP) in each color space as a predictor. The first set averaged NDOP across digits for each trial and used the averages and N as predictors in two-factor regressions for each color space. The second set consisted of four- and five-factor regressions on the four- and five-digit task-fields, respectively, in which NDOP for each digit was represented separately.

Proportion of discriminable outline pixels. The third modelling approach took into consideration the differences between digits in the number of outline pixels which they possessed. Presumably, the effect of a non-discriminable outline pixel is diminished for digits which possess a greater number of outline pixels. Therefore, the proportion of discriminable outline pixels (PDOP) was computed for each digit in each trial and was used in a set of regressions which were identical with those used for analysing NDOP.

Summary of results for outline model. The R^2s from all of the analyses which utilized the outline model are shown in Table 9. Nearly all R^2s are very low, indicating a major flaw in this model of the background colors. The R^2s for the two-factor models (which included N as a predictor) are barely larger than the R^2 obtained when a single-factor regression using N alone is performed on ΔRS, i.e., .056. The four-factor ΔE_{L*a*b*} and five-factor ΔE_{L*u*v*} and ΔE_{L*a*b} models achieved notably better R^2s than any others but even these do not compare favorably with those obtained previously with the two-degree model. Furthermore, it is unclear why L*u*v* and L*a*b* should perform so differently when four vs. five digits are displayed.

Part of the explanation for the apparent failure of the outline model may be that it does not account for the contributions of secondary pixels. For example, if all of the pixels surrounding the numeral "one" have the same color as the numeral, the outline model predicts minimum RS. Yet, if the outline pixels are themselves outlined by highly contrasting pixels, the numeral will nonetheless be legible and RS may not be affected appreciably.

Another problem with the particular approaches which were tested is that they assume a simple monotonically increasing relationship between the independent and dependent measures. Yet, it is easy to construct examples

Table 9. R^2s for Outline Model Regressions

	Two-factor	Four-factor	Five-factor
ΔE:			
Tri	.064	.062	.079
L*u*v*	.058	.056	.224
L*a*b*	.057	.118	.215
Wab	.075	.070	.094
NDOP:			
Tri	.063	.010	.009
L*u*v*	.060	.005	.008
L*a*b*	.063	.010	.009
Wab	.062	.013	.014
PDOP:			
Tri	.063	.009	.009
L*u*v*	.060	.004	.008
L*a*b*	.062	.009	.009
Wab	.062	.013	.013

in which a few nondiscriminable outline pixels, placed in critical positions, might be expected to produce the greatest decrements in RS. Thus, although this is not a fault inherent in the outline model, the analyses were insensitive to effects associated with the positions of the outline pixels.

Color Variation Metrics

Although it seemed plausible that additional work might improve the R^2s for the outline model, it was also clear this would require furthering its complexity, which was already computationally imposing. Since the simpler two-degree model had produced notably better R^2s, it appeared unlikely that any reasonable version of the outline model would prove superior. Therefore, it was abandoned and efforts were made to develop a new refinement of the two-degree model.

Passing consideration was given to the idea of decreasing the size and/or changing the shape of the averaging area. Taken to the smaller extreme this approach would, of course, have reproduced the outline model. It is possible that manipulation of the averaging area would have uncovered a demonstrably better geometry. However, this probably would have been very time-consuming and did not seem apt to produce substantial increases in R^2. The fact that the two-degree average produced very respectable

results when compared with Lippert's (1983) R^2s suggests that it is already a near-optimal representation of global color contrast. Therefore, it seemed probable that adding new information regarding color contrast would be more fruitful than attempting to "fine-tune" the two-degree radius average.

It has already been pointed out that the primary difference between the backgrounds used by Lippert (1983) vs. those used in the present study is their uniformity. Lippert's (1983) background pixels were monocolored whereas the present study's were multicolored. One might expect that increasing variability in the backgrounds' colors would lead to decreasing RS and, therefore, a metric which represents this variability should substantially improve the regression equations' R^2s. The second major effort to refine the two-degree model explored this possibility.

Definitions of variation metrics. Several metrics which represent the variability of the background pixels' colors were proposed. Because they were intended to serve as parameters in a two-degree radius model, they all restrict consideration to pixels within the two-degree averaging area. The metrics are named and defined as follows:

X: The mean distance in color space of the pixels from their colorimetric center of mass, i.e., the average coordinates

for the two-degree radius.

S: The standard deviation of the distances.

S^2: The variance of the distances. (This is simply a second-order version of S.)

CV: The coefficient of variation for the distances, i.e., S/X. Unlike X, S, and S^2, this is a unitless measure of variability.

CVR: A new and unitless metric which was named "color variation ratio". It is defined:

$$CVR = (U - 1) / (T - 1),$$

where U is the number of unique colors among the pixels within an arbitrary (two-degree, in the present case) radius and T is the total number of pixels within the radius. If all pixels have the same color, then U = 1 and CVR = 0. If all pixels are uniquely colored, then U = T and CVR = 1. The CVR is undefined for T = 1. A unique color is defined as one having a unique bit setting. This is equivalent to having unique tristimulus values, L*u*v* coordinates, etc. and, thus, the CVR is independent of the color space used to represent global color contrast. (For the other variation metrics, it only makes sense to compute them within the color space which is being used to represent global contrast.) The definition does not assume that the colors are unique perceptually.

Tests of variation metrics. The varition metrics were evaluated initially by performing two sets of regressions.

The first set consisted of a seven-factor regression for each color space, using N, ΔE, and all five variation metrics as parameters. Of course, the variation metrics are intercorrelated, but the resulting R^2s provided referents against which simpler models could be compared. The second set of regressions tested only one variation metric at a time, combining each with N and ΔE in each color space. The R^2s produced by the seven- and three-factor ΔE models are shown in Table 10 along with the R^2s from the original two-factor ΔE models, for purposes of comparison.

It is evident in Table 10 that S and CVR are the best single predictors. In all color spaces, they provided larger R^2s than did the other variation metrics and their averages across the color spaces are almost double the average for the original two-factor ΔE models. The average increase in R^2 for S and CVR was approximately .247, which is 73% of the average increase for the seven-factor models, i.e., .338.

The next step in evaluating the variation metrics consisted of computing the correlations between each metric and distance in each color space. Correlations with ΔE are shown in Table 11. Correlations with squared distance along each axis are shown in Table 12.

Examination of Tables 11 and 12 reveals that CVR is

Table 10. R^2S for Color Variation Metrics

Seven-factor ΔE regressions:

	Tri	L*u*v*	L*a*b*	Wab	Avg. R^2
	.587	.615	.585	.674	.615

Three-factor ΔE regressions:

	Tri	L*u*v*	L*a*b*	Wab	Avg. R^2
X	.370	.320	.351	.430	.368
S	.506	.484	.486	.601	.519
S^2	.424	.420	.441	.495	.445
CV	.261	.271	.284	.292	.277
CVR	.495	.543	.522	.557	.529
Avg. R^2:	.411	.408	.417	.475	.428

Original two-factor ΔE regressions:

Tri	L*u*v*	L*a*b*	Wab	Avg. R^2
.261	.271	.284	.292	.277

Table 11. Correlations Between Color Variation Metrics and ΔE in Each Color Space

	X	S	S^2	CV	CVR
Tri	.409	.263	.361	.052	-.095
L*u*v*	-.261	.279	.439	.490**	.058
L*a*b*	-.519**	.087	.172	.639**	-.134
Wab	.465**	.381	.468**	.055	-.016
Avg. |r|:	.413	.252	.360	.309	.076

**statistically significant using alpha = .05.

Table 12. Correlations Between Color Variation Metrics and Squared Distances for Color Spaces' Axes

	X	S	S^2	CV	CVR
$(\Delta X)^2$	.584**	.466**	.512**	-.053	.001
$(\Delta Y)^2$	.617**	.406	.415	-.124	-.061
$(\Delta Z)^2$	.506**	.113	.138	-.157	-.270
$(\Delta L^*)^2$	-.602**	-.348	-.211	.507**	-.400
$(\Delta u^*)^2$	.170	.672**	.814**	.212	.375
$(\Delta v^*)^2$	.151	.454**	.447**	.094	.332
$(\Delta L^*)^2$	-.688**	-.305	-.238	.588**	-.400
$(\Delta a^*)^2$	.173	.446**	.459**	.052	.293
$(\Delta b^*)^2$	.138	.519**	.559**	.155	.328
$(\Delta W)^2$	.586**	.350	.356	-.109	-.079
$(\Delta a)^2$	.252	.671**	.812**	.192	.349
$(\Delta b)^2$	.599**	.496**	.376	-.160	.183
Avg. |r|:	.422	.437	.445	.200	.256

** statistically significant using alpha = .05.

clearly the least troublesome variation metric with regard to covariance with HUD/background distance parameters. Its average absolute correlation with ΔE is much smaller than those of the other metrics and its average absolute correlation with the squared distances is almost as small as CV's, which is the poorest predictor in Table 10. Furthermore, unlike the other metrics, none of CVR's correlations are statistically significant. For these correlations, the issue of statistical significance is important because it is worthwhile to know whether the covariance which is evident in the present sample exists in the (infinite) population of potential stimuli. On this basis, CVR is unquestionably the best metric because it is least likely to covary with global color contrast parameters in one's regression equations. The CVR is also the best metric for purposes of the present analysis because it combines superior predictive power with relatively low covariance for the present experiment's HUDBACKSs. Therefore, it was selected for inclusion with the original two-degree model of color contrast, yielding the new "two-degree model of color variation."

Two-degree Model of Color Variation

Three-factor ΔE regressions. The R^2s for the three-factor ΔE models have already been presented but the regressions involving CVR are shown in more detail in Table

Table 13. Two-Degree Model of Color Variation:

Three-Factor ΔE Regressions

Parameter	Regression coefficient	p	R^2
Intercept	3.126380×10^{-3}	< .01	.495
N	-2.745800×10^{-4}	< .01	
CVR	-1.084280×10^{-3}	< .01	
ΔE_{Tri}	6.218733×10^{-6}	< .01	
Intercept	2.973990×10^{-3}	< .01	.543
N	-2.745800×10^{-4}	< .01	
CVR	-1.116477×10^{-3}	< .01	
ΔE_{L*u*v*}	5.867772×10^{-6}	< .01	
Intercept	2.927130×10^{-3}	< .01	.522
N	-2.745800×10^{-4}	< .01	
CVR	-1.092970×10^{-3}	< .01	
ΔE_{L*a*b*}	6.902558×10^{-6}	< .01	
Intercept	3.079510×10^{-3}	< .01	.557
N	-2.745800×10^{-4}	< .01	
CVR	-1.150330×10^{-3}	< .01	
ΔE_{Wab}	3.307664×10^{-5}	< .01	

13. All parameters in all color spaces were significant, $p < .01$ in all cases, and the sign of CVR's slopes was negative, as predicted. The covariance between CVR and ΔE is evident in Table 13, wherein the estimate of CVR's slope varies slightly across the color spaces. Also, comparison with Table 2 shows that adding CVR to the regression equations caused the intercepts and slopes for ΔE to shift a bit.

Five-factor regressions. The next logicl step was to test regression equations which included CVR and permitted the color spaces' axes to be rescaled. These regressions took essentially the same form as Equation 1 but with CVR, as well as N, subtracted from the right-hand side. The slope for CVR was estimated in advance via a two-factor regression which used N and CVR as parameters.

The results for the five-factor regressions are summarized in Table 14. The estimated slopes for Tri's and Y and Z axes and L*a*b*'s b axis did not differ significantly from zero, $p > .10$ in all cases, but all other parameters were significant, $p < .02$ in all cases. This appears to be a slight improvement over Table 3. Comparisons of the five-factor R^2s with their three-factor counterparts showed no significant differences for Tri or Wab, $p > .18$ in both cases, and showed the five-factor R^2s for L*u*v* and L*a*b* to be significantly greater, $p < .01$

Table 14. Two-Degree Model of Color Variation: Five-Factor Regressions

Parameter	Regression coefficient	p	R^2
Intercept	1.180573×10^{-05}	< .01	.469
N	$-2.745800 \times 10^{-04}$	< .01	
CVR	$-1.152040 \times 10^{-03}$	< .01	
$(\Delta X)^2$	6.331391×10^{-10}	< .01	
$(\Delta Y)^2$	4.150579×10^{-11}	= .80	
$(\Delta Z)^2$	$-2.523097 \times 10^{-10}$	= .10	
Intercept	1.012723×10^{-05}	< .01	.639
N	$-2.745800 \times 10^{-04}$	< .01	
CVR	$-1.152040 \times 10^{-03}$	< .01	
$(\Delta u^*)^2$	9.274627×10^{-10}	< .01	
$(\Delta u^*)^2$	1.936028×10^{-10}	< .01	
$(\Delta v^*)^2$	7.927503×10^{-11}	= .02	
Intercept	9.997982×10^{-06}	< .01	.634
N	$-2.745800 \times 10^{-04}$	< .01	
CVR	$-1.152040 \times 10^{-03}$	< .01	
$(\Delta L^*)^2$	9.773232×10^{-10}	< .01	
$(\Delta a^*)^2$	4.069501×10^{-10}	< .01	
$(\Delta b^*)^2$	8.710482×10^{-12}	= .86	

Table 14 (continued)

Intercept	1.070769×10^{-05}	< .01	.539
N	$-2.745800 \times 10^{-04}$	< .01	
CVR	$-1.152040 \times 10^{-03}$	< .01	
$(\Delta W)^2$	4.344408×10^{-09}	< .01	
$(\Delta a)^2$	9.849166×10^{-09}	< .01	
$(\Delta b)^2$	2.047066×10^{-08}	< .01	

Table 16. Two-Degree Model of Color Variation vs. Lippert (1983) Models: Predicting Lippert's (1983) Results

Color space	L model	R^2	P model	R^2	p
Tri	4F	.597	3F- ΔE	.289	< .01
L*u*v*	4F	.656	5F	.362	< .01
L*a*b*	4F	.619	5F	.320	< .01
Wab	2F- ΔE	.425	3F- ΔE	.405	= .05

L = Lippert (1983)

P = Present study

4F = Four-factor

5F = Five-factor

2F- ΔE = Two-factor ΔE

3F- ΔE = Three-factor ΔE

L*a*b*, and Tri were significant, $p < .01$ in all cases, and were in favor of a given study's models predicting that study's results. The results for Wab were also essentially the same as before. For predicting the present study's results, either experiment's Wab coefficients served equally well and, for predicting Lippert's (1983) results, the comparison was barely significant. It is noteworthy that the former outcome resulted from a remarkable improvement in R^2 which resulted from adding CVR to the Lippert (1983) Wab model and which the other color spaces did not enjoy equally. (Compare the Lippert (1983) R^2s in Tables 7 vs. 15.)

Comparison of the R^2s for the present study's models in Tables 8 vs. 16 suggests that the new optimum models did not predict the results from Lippert (1983) any better than did the original optimum models. However, examination of Tables 15 and 16 indicates that the new optimum models predict the present study's results as well as the optimum Lippert (1983) models predict Lippert's (1983) results. Thus, it is unlikely that substantial improvements in these R^2s for the present study can be achieved.

Nonlinear regressions. A final check was made to assure that the coefficients for the five-factor models were optimal. It was possible that the estimated intercepts and slopes for the color spaces' axes had been biased by the

subtraction of a covarying term (i.e., CVR) from the dependent measure prior to the regressions. Recall that the slope for CVR had been estimated independently of squared distance along the axes, while the axes' slopes had been estimated given the slope for CVR. Thus, it was desirable to check the coefficients in a regression which would estimate all parameters simultaneously.

The equations to which a fit was desired took a form similar to Equation 2, e.g.:

$$[x_0 + x_1(dL^*)^2 + x_2(du^*)^2 + x_3(dv^*)^2]^{1/2} + x_4N + x_5CVR = RS \quad (4)$$

where x_5 is the regression coefficient for CVR. It has already been pointed out that equations such as this cannot be fitted using linear regression methods. However, solutions can sometimes be achieved using nonlinear search techniques which perform least-squares minimization, particularly if one has some general notion as to the location of the solution. Although the coefficients in Table 14 might not be optimal, it is likely that they are close and, therefore, they provided suitable starting points for searches in the color spaces.

The searches utilized the multivariate secant method, known also as DUD or the method of false position, as implemented in the Statistical Analysis System (SAS

Institute, 1982). The multivariate secant method is similar to the classical Gauss-Newton procedure but it estimates the derivatives of the regression parameters from its own sequence of iterations.

A set of searches was performed for each color space. The first started at the location specified by the coefficients in Table 14 while subsequent searches started from slightly different locations. Convergence was achieved in all cases without evidence of computational error or difficulties. The optimized coefficients resulting from the searches did not differ appreciably from those in Table 14 in any case nor did the R^2s improve notably (The average increase in R^2 was only .006.) Furthermore, the searches perturbed the estimates of N's slope, which should have remained invariant. Therefore, the coefficients in Table 14 were deemed satisfactory.

DISCUSSION AND CONCLUSIONS

It is evident that color contrast metrics which utilize Euclidean distance in color space can be related meaningfully to human visual task performance. Although the outline model was largely unsuccessful, all of the regressions for the other major models produced respectable R^2s.

The differences among the R^2s as a function of color space and model complexity are instructive with regard to the relative practical merits of the color spaces which were studied. The present study's ΔE regressions indicate that none of the uniform color spaces, as they are defined presently by the CIE 1976 formulations, provide any substantial improvements over the 1931 CIE tristimulus space in one's ability to predict performance as a function of color distance. This finding is difficult to reconcile completely with the results from the ΔE regressions on Lippert's (1983) data. Perhaps the most reasonable conclusion is that any benefits which are associated with using the uniform color spaces in their present (CIE 1976) form are situation-specific and, at best, do not appear to be especially great.

The present results also indicate that a simple linear rescaling of the L*u*v* and L*a*b* axes yields substantial

and equivalent improvements in their predictive utility, but does not benefit the 1931 CIE and Wab spaces. Here again, however, the finding for 1931 CIE space does not agree totally with that based on Lippert's (1983) data, although Lippert achieved substantial improvement in prediction of visual task performance by rescaling the L*u*v* axes.

The aforementioned results for L*u*v* and L*a*b* contribute to a growing body of evidence (i.e., Costanza, 1981; Pointer, 1981) which suggests that these two spaces are equally useful but are not perceptually uniform. For both spaces, all rescaled-axis regressions which are discussed herein yielded axis slopes which differed considerably. The same argument can be made in the case of Wab space, although the findings suggest that something other than a linear rescaling will be required to produce any improvement over the original transformation. For L*u*v* and L*a*b*, however, it is clear that merely re-weighting their axes should produce substantial and consistent benefits. Furthermore, it appears that this adjustment should render these two spaces superior to the Wab and 1931 CIE spaces.

The present results also have interesting implications regarding color selection for HUDs. The red, yellow-green, and achromatic HUDs were intended to simulate chromaticities which are available using contemporary HUD technology.

Clearly, the red HUD yielded the best overall performance while the other two yielded the worst. As mentioned previously, though, the applicability of these findings is dependent upon the representativeness of the color contrasts which were presented. Appendices 3 - 6 are intended to assist the reader in evaluating the stimuli which were used in the present study.

Appendix 3 tabulates the backgrounds' colorimetric characterizations, as used in the present analyses. Appendix 4 shows the locations of the two-degree averages and the five HUD chromaticities on the 1931 CIE chromaticity diagram. The triangle in Appendix 4 which has the red, green, and blue HUD chromaticities (which were each produced by a single gun) as its apices indicates the range of chromaticities which the monitor was capable of producing. Appendix 5 summarizes the characterizations of Renndorf (1956) for various natural and man-made objects, and Appendix 6 shows the locations of these chromaticities on the 1931 CIE chromaticity diagram.

Comparison of the chromaticity coordinates in these tables and figures does not reveal any obvious and major differences. However, it is possible for luminance contrasts to exist in operational environments which exceed those used in the present experiment. Furthermore, it is possible that effects associated with retinal illuminance

and ambient illumination mediate those obtained in the present study. Finally, it is not certain whether the differences in RS which were obtained as a function of HUD color exist or are meaningful in applied settings. Therefore, although it seems likely that a red HUD is optimal, this finding must be regarded as tentative.

It is interesting that the red and blue HUDs, which yielded the best overall performance, can be seen in Appendix 4 to lie farthest from the two-degree radius averages. However, this observation must be tempered by the realization that the 1931 CIE diagram is not perceptually uniform. Also, it should be borne in mind that the points which represent the backgrounds are averages and, hence, a larger portion of the chromaticity diagram was sampled than might be inferred from the plot.

Perhaps the most important result of the present research is the finding that a relatively simple approach to modelling spatially complex color contrasts accounted for substantial portions of the sample variation. This has great practical significance. The computation of tristimulus values for a two-degree averaging area and subsequent transformation to a uniform color space can be accomplished quickly and easily, even without a computer or digital image processor. As the color contrast model shows, this alone may be expected to provide large R^2s if the

rescaled L*u*v* or L*a*b* spaces are used. Even better results can be had from these two spaces if the CVR is also computed. This is an extremely straightforward procedure if the image is in digital form.

REFERENCES

Bloomfield, J. R. Visual search with embedded targets: color and texture differences. Human Factors 1979, 21, 317-330.

Cohen, J. and Friden, T. P. The Euclidean nature of color space, Bulletin of the Psychonomic Society, 1975, 5, 159-161.

Costanza, E. B. An evaluation of a method to determine suprathreshold color contrast on CRT displays. Unpublished Master's Thesis, Virginia Polytechnic Institute and State University, 1981.

Lippert, T. M. Color contrast effects for a simulated CRT head-up display. Unpublished Master's Thesis, Virginia Polytechnic Institute and State University, 1983.

Lippert, T. M., Farley, W. W., Post, D. L., and Snyder, H. L. Color contrast effects on visual performance. Digest of the Society for Information Display, 1983, 170-171.

Lippert, T. M. and Snyder, H. L. Unitary suprathreshold color-difference metrics of legibility for CRT raster imagery. Virginia Polytechnic Institute and State University Technical Report HFL/ONR-86-3, 1986.

McNemar, Q. Psychological statistics. New York: Wiley, 1955, 148.

Meister, D. and Sullivan, D. J. Guide to human engineering design for visual displays. Washington: ONR Report under Contract N00014-68-C-0278, 1969.

Penndorf, R. Luminous and spectral reflectance as well as colors of natural objects. Bedford, Massachusetts: Geophysics Research Directorate, Geophysical Research Paper No. 44, February, 1956.

Pointer M. R. A comparison of the CIE 1976 colour spaces. Color Research and Application, 1981, 6, 108-118.

Post, D. L., Costanza, E. B., and Lippert, T. M. Expressions of color contrast as equivalent achromatic contrast. Proceedings of the Human Factors Society 26th Annual Meeting, 1982, 581-585.

Rizy, E. F. Dichroic filter specification for color additive displays. RADC Technical Report No. TR-67-513, 1967.

SAS Institute. SAS user's guide: statistics. Cary, North Carolina: SAS Institute Inc., 1982.

APPENDIX 1

Average Photometric Measurements for HUD Colors

	x	y	Y (cd/m^2)
Yellow-green	.3757	.4918	39.01
Red	.6088	.3429	40.01
Achromatic	.3112	.3282	38.40
Blue	.1535	.0693	42.02
Green	.3038	.5625	39.59

APPENDIX 2

Descriptions of Photographs Used as Backgrounds

1. A harbour, photographed from an elevated site approximately one mile away. Water is light green, nearby buildings brown, gray, and tan with interspersed greenery, and large brown mountain overlooking scene in background.
2. Open sea with horizon and pale blue sky with a few clouds visible. Most of the water is dark blue with black ripples. Shallow area contains light blue water.
3. Bare, cracked, light brown ground with gray and brown mountain and light blue sky in background.
4. A concrete spillway in a wooded lake area with barren trees. Water dark blue with white foam near shore and spillway.
5. A gray U.S. destroyer, side view, photographed such that its length slightly exceeded the photograph's width. Water is very dark blue. Sky is pale blue with a few clouds.
6. A flat two-lane country road, photographed head-on from an overpass. Two nearby automobiles visible. Lush green trees on either side, yellow grass on enbankment, and grayish-blue asphalt.
7. View from overpass of same road in opposite direction. Road climbs uphill and is white with distant automobile visible. Green trees and grass on either side.
8. False-colored Landsat photograph of city and nearby mountain range. Mostly reds and blues with some greens, whites,

and browns.

9. Aerial photograph of a lake with Pacific ocean nearby. Ground is dark green, brown, and gray, lake is gray, ocean is bluish-gray, and scattered clouds just below aircraft are white.
10. Aerial photograph with heavily wooded terrain barely visible through blue haze. Two large white clouds at top and on right.

APPENDIX 3

Characterization of Two-Degree Averages for Backgrounds

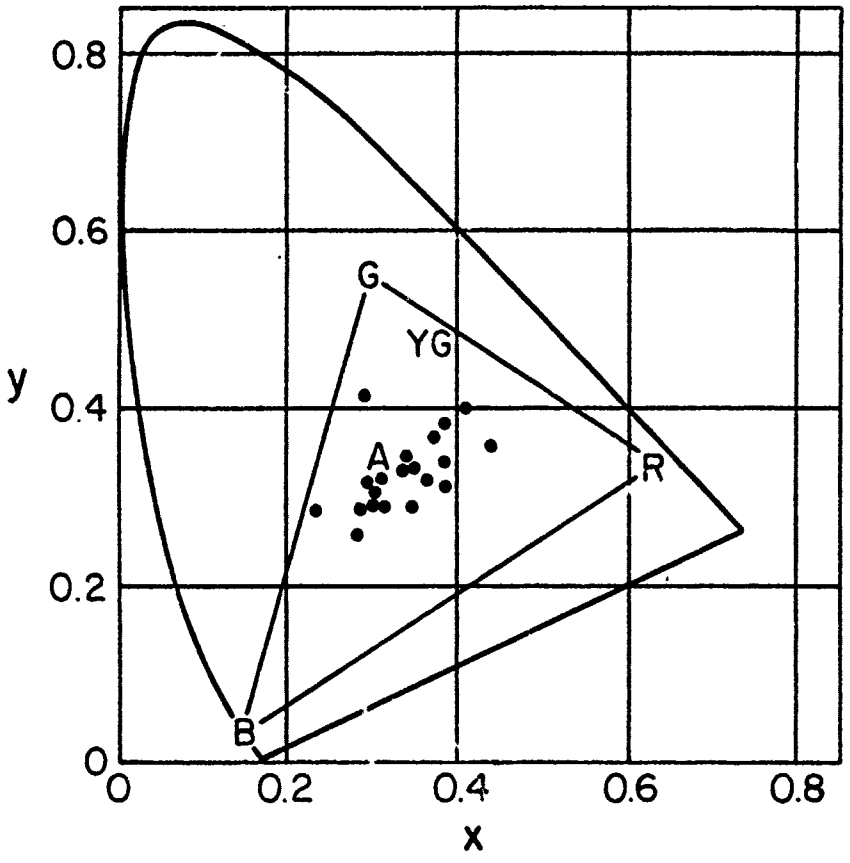
B	Photo	x	y	Y (cd/m^2)	CVR	DW	Purity
1	1	.3804	.3613	1.331	.7013	584	.24
2	1	.3990	.3452	1.673	.7583	596	.25
3	2	.2773	.2565	1.469	.5235	466	.26
4	2	.2385	.2939	12.585	.4816	487	.34
5	3	.3934	.3759	19.634	.9357	580	.36
6	3	.3686	.3281	5.649	.8690	620	.11
7	4	.3391	.3389	27.587	.8929	582	.05
8	4	.2914	.2869	20.365	.7257	475	.18
9	5	.3163	.3316	34.097	.6216	495	.05
10	5	.3069	.3052	27.371	.7076	470	.10
11	6	.3416	.3489	110.290	.5217	571	.09
12	6	.4174	.4041	70.011	.9734	580	.46
13	7	.3042	.3606	54.893	.9547	507	.09
14	7	.2881	.4214	10.322	.9582	524	.18
15	8	.4364	.3634	67.477	.9852	593	.41
16	8	.3542	.3407	94.166	.9861	585	.10
17	9	.3482	.2972	1.056	.6150	-519	.16
18	9	.2997	.2857	17.698	8650	469	.16
19	10	.3787	.3186	0.863	.2730	-495	.15
20	10	.3013	.3049	53.809	.3459	478	.12

B = Background

DW = Dominant wavelength

APPENDIX 4

Chromaticity Plot for Background Images and HUD Colors



APPENDIX 5

Colors in the Natural Environment

	x	y	Y(%)	Dominant wavelength	Purity
Inland water	.269	.289	5	481	.31
Snow	.340	.346	77	481	.03
Ice	.351	.354	75	579	.02
Limestone clay	.377	.376	63	579	.18
Bare mountaintops	.399	.387	24	582	.29
Dry sand	.399	.387	24	582	.29
Wet clay soil	.382	.373	9	583	.18
Bare dry ground	.382	.373	9	583	.18
Black earth, sand, loam	.377	.369	3	583	.15
Coniferous forests, winter	.381	.396	3	574	.25
Coniferous forests, summer	.397	.410	8	576	.36
Deciduous forests, fall	.451	.399	15	586	.50
Deciduous forests, summer	.394	.432	10	572	.43
Lush grass	.394	.432	10	572	.43
Dry meadow grass	.397	.410	8	576	.36

Appendix 5 (continued)

Ripe field crops	.451	.399	15	586	.50
Earth roads	.377	.369	3	583	.15
Blacktop roads	.382	.373	9	583	.18

APPENDIX 6

Chromaticity Plot of Colors in the Natural Environment